How Libraries can Benefit from Linked Data

Sandra Gonzalez

Marshall School of Business, University of Southern California

The worldwide web has afforded us an abundance of information to access. How then do we manage to filter through endless web pages and records to locate what is needed? Much of the work to define and describe data is completed behind the scenes. Linked Data (LD) is one method of encoding and publishing data on the web (C. Muglia, personal communication, May 26, 2021). However, LD cannot be understood without first comprehending the Semantic Web (SW) and Resource Description Framework (RDF). One way of structuring all this data on the web is through what is called the Semantic Web. SW can be explained as a type of web configuration used to read and recognize data quickly. Resource Description Framework (RDF) refers to a model used to process data interchangeably. Essentially it is a "sentence structure for metadata statements with three parts: a subject, a predicate, and an object" (C. Muglia, personal communication, May 26, 2021). "Linked data uses RDF as its foundation to ensure this clarity in meaning" (Taylor & Joudrey, 2018, p. 33). LD is the future of description and access in Libraries. It provides greater accuracy, improved the discovery of rare resources, and increases the visibility of library resources on the web.

Linked data delivers greater accuracy with easy use and increased access. "LD is machine-readable, has meaning explicitly designed, linked to other external data sets, and is linked to *from* external data sets" (Taylor & Joudrey, 2018, p. 33). It can share data from various materials and creates a type of automation. Both librarians and users stand to benefit from this automation. Sharing capability eliminates the additional leg work put in by catalogers and library administrators who often spend many hours creating metadata, integrating, and formatting different materials. Instead, catalogers can focus on the materials themselves.

Moreover, LD can be used to add more contextualized content in metadata, which means that searches could make distinctions between results (Taylor & Joudrey, 2018, p. 129), thus

improving discoverability on the web. When more description is added, more discoverability is possible. As a result, libraries can create a digital library filled with digital materials found in the physical library and more. Essentially, this method direct access to materials. Every action in linked data creates a domino effect. The easier it is for a user to search, discover, and access information, the more significant the interaction increases.

Linked data expands opportunities for discovery. With Uniform Resource Identifiers (URI), LD includes links to other UIR's so users can discover more items (Schilling, 2012). URI's help identify objects and abstract concepts. Essentially, Linked Data also needs URI to work effectively. URI's play a meaningful role in identifying resources one did not know they would need. With linked data, one can stumble upon additional, valuable, and rare information when searching for something specific. In terms of inclusivity, this aspect of LD can have a considerable, positive impact on the future of information retrieval.

Additionally, linked data could open doors for more inclusive content that users would not otherwise know to locate. "A goal of the Semantic Web is to increase technology's abilities to connect an even wider variety of related resources through automatic means" (Taylor & Joudrey, 2018, p. 33). User searches can be limited depending on their experience conducting research or using Boolean search terms. The idea of retrieving more resources outside of the scope of the initial search can support resources that are often overlooked and expand user knowledge beyond homogenous results.

Linked data has the potential to enhance the web using structured data (Schilling, 2021). Libraries currently do not have the capability of applying their materials on the web. Users can access the libraries' resources only through the library interface or catalog. This limits the libraries' reach. Schilling (2012) suggests that libraries leverage metadata as linked data to be

included in the search results obtained by users online. The linking of data online expands results tremendously. Serendipity is a term often referred to in explaining the results of linked data on the web. *Serendipity* is defined as events that happen by chance. As a user, linked data can present many serendipitous results on online searches. Perhaps a user is not narrowing or expanding their search results enough, but because of the inclusion of linked data through RDF and URI, the lack of Boolean search terms experience will not matter. The web will provide the results an individual is searching for without even knowing it. After all, "users cannot search for unknown information or resources" (Schilling, 2012, p. 5). The possibilities for discoverability are endless. Linked data will effectively blend two worlds: libraries and "non-library services" (Schilling, 2012, p. 5).

The internet has become a go-to place for everything. Some say that it has or will replace libraries, but libraries are evolving simultaneously as technology and the internet. The web is no match for libraries. On the contrary, libraries can use the web as a tool to enhance their presence. Knowledge institutions are constantly seeking ways to increase patron visits. More patrons mean increased budgets. Increased budgets mean more programs, services, and resources for patrons. Because linked data makes structured data on the web, so they are easily understood by computers, it can aid in the discovery of knowledge organizations. Thus, new library programs, collections, and events will be easily accessed by future patrons.

Moreover, LD could push libraries to design an open collection of resources on the web. With RDF and URI, resource descriptions will be made more widely available. The sharing of these descriptions will benefit librarians and catalogers. They will be able to pull this information and re-use it. It is a time-saving technique that will focus on developing collections, creating programs, or other essential tasks. RDF and URI will also impact their relevancy through Search

Engine Optimization (SEO), which is already widely used by many private companies and libraries with a budget. SEO aids in retrieving information when users are searching on the web, and SEO is somewhat related to online algorithms. LD is a considerable opportunity that libraries can use to their advantage.

Additionally, many librarians have discussed the use of a cloud-based system. LD is underutilized in libraries, mainly because they do not have the budget to transition from MARC-based systems to the creation of linked data (Schilling, 2012). Converting various schemas into LD is very complex. The transition will take much time and expertise grounded on metadata and conversions on the web (Schilling, 2012). So, librarians are coming up with ways to alleviate the complexities of attempting to standardize the schemas and, in its place, create a system that will be less complicated and within a library's budget. This is where a cloud-based system would come into play. It is inexpensive and more manageable, but it will not work unless it becomes the standard and adopted by all libraries.

Another way libraries are trying to make the switch is through the Bibliographic Framework Initiative (BIBFRAME). The Library of Congress (LOC) created this project as a replacement for MARC, which is the primary encoding standard for library-generated metadata (C. Muglia, personal communication, May 26, 2021). BIBFRAME is essentially a shortcut to link data through the Semantic Web and makes library data available online. As of the time of this writing, library records are not available online. For example, if a user were to conduct a Google search for a book, they would receive many results, but none corresponds to library records. With BIBFRAME, when a user searches for a book on Google, they will discover Amazon and Barnes and Noble links and links to various library records such as Libby or a

catalog record. BIBFRAME will also enhance the connection between libraries creating a more extensive network where information can be shared.

To sum up, there are many benefits to Linked Data that outweigh the challenges. Linked data stands to serve patrons and all knowledge institutions in several ways, including greater accuracy, the improved discovery of rare resources, and increased visibility of library resources on the web while also creating a more extensive web of connectivity among libraries. Projects such as BIBFRAME are already in place to pave the way for linked data to be fully functional. RDF and URI are significant factors of linked data and the Semantic web. The real challenge is getting libraries to agree to move forward with this standard.

References

Schilling, V. (2012). *Transforming library metadata into linked library data.* American Library

   Association. http://www.ala.org/alcts/resources/org/cat/research/linked-data

Taylor, A. G., & Joudrey, D. N. (2018) *The Organization of Information*. (4th Ed.). Library and

   Information Science Text Series. Westport, Connecticut: Libraries Unlimited.